OIPE

RAW SEQUENCE LISTING DATE: 10/16/2001
PATENT APPLICATION: US/09/966,422 TIME: 15:52:20

Input Set : A:\30534111.app
Output Set: N:\CRF3\10162001\I966422.raw

3 <110> APPLICANT: Feder, J. N.
4 Mintier, G.
5 Ramanathan, C. S.
6 Hawken, D. R.
7 Cacace, A.
8 Barber, L.
9 Kornacker, M. G.
11 <120> TITLE OF INVENTION: A NOVEL HUMAN G-PROTEIN COUPLED RECEPTOR, HGPRBMY6,
12 EXPRESSED HIGHLY IN SMALL INTESTINE
14 <130> FILE REFERENCE: D0040NP/3053-4119US3
C--> 16 <140> CURRENT APPLICATION NUMBER: US/09/966,422 01
17 <141> CURRENT FILING DATE: 2001-09-26
19 <150> PRIOR APPLICATION NUMBER: 60/235,602
20 <151> PRIOR FILING DATE: 2000-09-27
22 <150> PRIOR APPLICATION NUMBER: 60/306,604
23 <151> PRIOR FILING DATE: 2001-07-19
25 <150> PRIOR APPLICATION NUMBER: 60/315,412
26 <151> PRIOR FILING DATE: 2001-08-28
28 <160> NUMBER OF SEQ ID NOS: 81
30 <170> SOFTWARE: PatentIn Ver. 2.1
32 <210> SEQ ID NO: 1
33 <211> LENGTH: 1683
34 <212> TYPE: DNA
35 <213> ORGANISM: Homo sapiens
37 <400> SEQUENCE: 1
38 atggagactt attccttgtc tttgggtaat caatcagtgg tggaacctaa catagcaata 60
39 cagtcagcaa atttctcttc agaaaatgcg gtggggcctt caaatgttcg cttctctgtg 120
40 cagaaaggag ctagcagttc tctagtttct agttcaacat ttatacatac aaatgtggat 180
41 ggccttaacc cagatgcaca gactgagctt caggtcttgc ttaatatgac gaaaaattac 240
42 accaagacat gcggctttgt agtttatcaa aatgacaagc ttttccaatc aaaaactttt 300
43 acagctaaat cggattttag tcaaaaaatt atctcaagca aaactgatga aaatgagcaa 360
44 gatcagagtg cttctgttga catggtcttt agtccaaagt acaaccaaaa agaatttcaa 420
45 ctctattcct atgcctgtgt ctattggaat ttgtcagcga aggactggga cacatatggc 480
46 tgtcaaaaag acaagggcac tgatggattc ctgcgctgcc gctgcaacca tactactaat 540
47 tttgctgtat taatgacttt caaaaaggat tatcaatatc ccaaatcact tgacatatta 600
48 tccaacgttg gatgtgcact gtctgttact ggtctggctc tcacagttat atttcagatt 660
49 gtcaccagga aagtcagaaa aacctcagta acctgggttt tggtcaatct gtgcatatca 720
50 atgttgattt tcaacctcct ctttgtgttt ggaattgaaa actccaataa gaacttgcag 780
51 acaagtgatg gtgacatcaa taatattgac tttgacaata atgacatacc caggacagac 840
52 accattaaca tcccgaatcc catgtgcact gcgattgccg ccttactgca ctatttctg 900
53 ttagtgacat ttacctggaa cgcactcagc gctgcacagc tctattacct tctaataagg 960
54 accatgaagc ctcttcctcg gcatttcatt cttttcatct cattaattgg atggggagtc 1020
55 ccagctatag tagtggctat aacagtggga gttatttatt ctcagaatgg aaataatcca 1080
56 cagtgggaat tagactaccg gcaagagaaa atctgctggc tggcaattcc agaacccaat 1140
57 ggtgttataa aaagtccgct gttgtggtca ttcatcgtac ctgtaaccat tatcctcatc 1200
58 agcaatgttg ttatgtttat tacaatctcg atcaaagtgc tgtggaagaa taaccagaac 1260
59 ctgacaagca caaaaaaagt ttcatccatg aagaagattg ttagcacatt atctgttgca 1320

Does Not Comply
Corrected Diskette Needed

Errored
See Error Summary Sheet
See page 6 of 7A

Raw Sequence Listing Error Summary

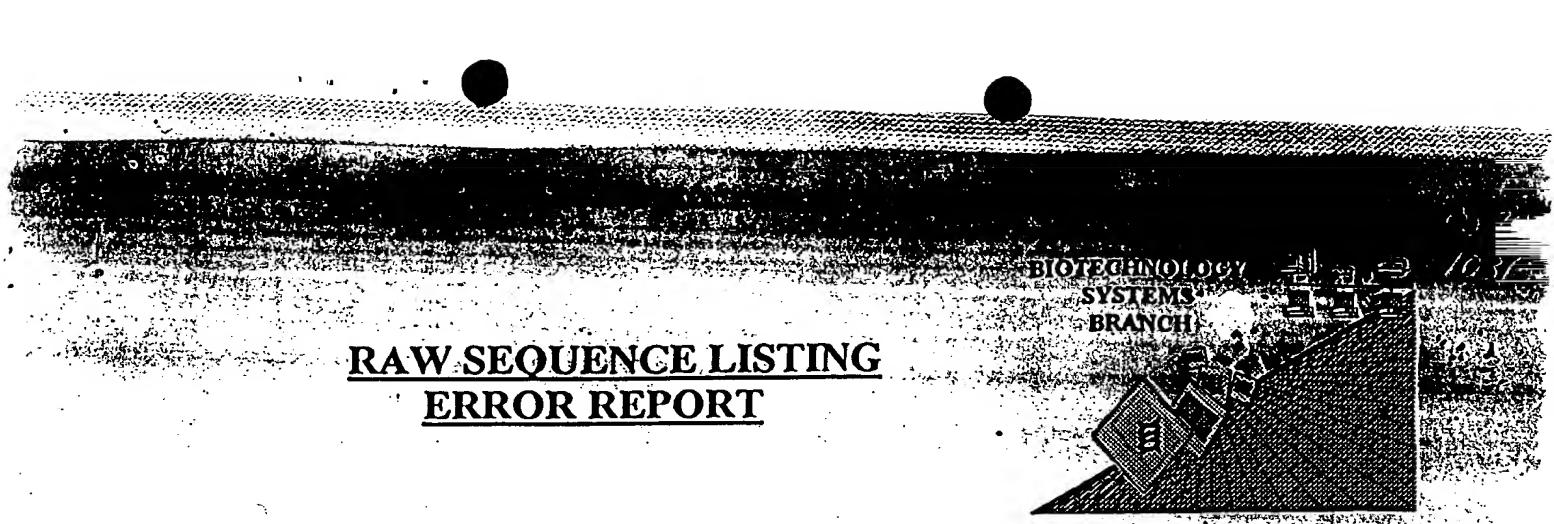
ERROR DETECTED	SUGGESTED CORRECTION

SERIAL NUMBER: 09/966422

ATTN: NEW RULES CASES: PLEASE DISREGARD ENGLISH "ALPHA" HEADERS, WHICH WERE INSERTED BY

1____ Wrapped Nucleics Wrapped Aminos	The number/text at the end of each line "wrapped" down to the next line. This may occur if your fi was retrieved in a word processor after creating it. Please adjust your right margin to .3; this will prevent "wrapping."
2____ Invalid Line Length	The rules require that a line not exceed 72 characters in length. This includes white spaces.
3____ Misaligned Amino Numbering	The numbering under each 5th amino acid is misaligned. Do not use tab codes between numbers; use space characters, instead.
4____ Non-ASCII	The submitted file was not saved in ASCII(DOS) text, as required by the Sequence Rules. Please ensure your subsequent submission is saved in ASCII text.
5____ Variable Length.	Sequence(s)____ contain n's or Xaa's representing more than one residue. Per Sequence Rules, each n or Xaa can only represent a single residue. Please present the maximum number of each residue having variable length and indicate in the <220>-<223> section that some may be missing.
6____ PatentIn 2.0 "bug"	A "bug" in PatentIn version 2.0 has caused the <220>-<223> section to be missing from amino acid sequences(s) _____. Normally, PatentIn would automatically generate this section from the previously coded nucleic acid sequence. Please manually copy the relevant <220>-<223> section to the subsequent amino acid sequence. This applies to the mandatory <220>-<223> sections for Artificial or Unknown sequences.
7____ Skipped Sequences (OLD RULES)	Sequence(s) ____ missing. If intentional, please insert the following lines for each skipped sequenc (2) INFORMATION FOR SEQ ID NO:X: (insert SEQ ID NO where "X" is shown) (i) SEQUENCE CHARACTERISTICS: (Do not insert any subheadings under this heading) (xi) SEQUENCE DESCRIPTION:SEQ ID NO:X: (insert SEQ ID NO where "X" is shown) This sequence is intentionally skipped Please also adjust the "(ii) NUMBER OF SEQUENCES:" response to include the skipped sequences
8____ Skipped Sequences (NEW RULES)	Sequence(s) ____ missing. If intentional, please insert the following lines for each skipped sequen <210> sequence id number <400> sequence id number 000
9____ Use of n's or Xaa's (NEW RULES)	Use of n's and/or Xaa's have been detected in the Sequence Listing. Per 1.823 of Sequence Rules, use of <220>-<223> is MANDATORY if n's or Xaa's are present. In <220> to <223> section, please explain location of n or Xaa, and which residue n or Xaa represent
10____ Invalid <213> Response	Per 1.823 of Sequence Rules, the only valid <213> responses are: Unknown, Artificial Sequence, or scientific name (Genus/species). <220>-<223> section is required when <213> response is Unknown or is Artificial Sequence
11 ✓ Use of <220>	Sequence(s) _____ missing the <220> "Feature" and associated numeric identifiers and responses. Use of <220> to <223> is MANDATORY if <213> "Organism" response is "Artificial Sequence" or "Unknown." Please explain source of genetic material in <220> to <223> section. (See "Federal Register," 06/01/1998, Vol. 63, No. 104, pp. 29631-32) (Sec. 1.823 of Sequence Rules
12____ PatentIn 2.0 "bug"	Please do not use "Copy to Disk" function of PatentIn version 2.0. This causes a corrupted file, resulting in missing mandatory numeric identifiers and responses (as indicated on raw sequence listing). Instead, please use "File Manager" or any other manual means to copy file to floppy disk.
13____ Misuse of n	n can only be used to represent a single nucleotide in a nucleic acid sequence. N is not used to represe any value not specifically a nucleotide.

AMC/MH – Biotechnology Systems Branch – 08/21/2001



RAW SEQUENCE LISTING ERROR REPORT

The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) detected errors when processing the following computer readable form:

Application Serial Number: 09/966422

Source: OIPE

Date Processed by STIC: 10/16/01

THE ATTACHED PRINTOUT EXPLAINS DETECTED ERRORS.
PLEASE FORWARD THIS INFORMATION TO THE APPLICANT BY EITHER:

1) **INCLUDING A COPY OF THIS PRINTOUT IN YOUR NEXT COMMUNICATION TO THE APPLICANT, WITH A NOTICE TO COMPLY or,**
2) **TELEPHONING APPLICANT AND FAXING A COPY OF THIS PRINTOUT, WITH A NOTICE TO COMPLY**

FOR CRF SUBMISSION QUESTIONS, PLEASE CONTACT MARK SPENCER, 703-308-4212.

FOR SEQUENCE RULES INTERPRETATION, PLEASE CONTACT ROBERT WAX, 703-308-4216.
PATENTIN 2.1 e-mail help: patin21help@uspto.gov or phone 703-306-4119 (R. Wax)
PATENTIN 3.0 e-mail help: patin3help@uspto.gov or phone 703-306-4119 (R. Wax)

TO REDUCE ERRORED SEQUENCE LISTINGS, PLEASE USE THE CHECKER VERSION 3.0 PROGRAM, ACCESSIBLE THROUGH THE U.S. PATENT AND TRADEMARK OFFICE WEBSITE. SEE BELOW:

Checker Version 3.0

The Checker Version 3.0 application is a state-of the-art Windows based software program employing a logical and intuitive user-interface to check whether a sequence listing is in compliance with format and content rules. Checker Version 3.0 works for sequence listings generated for the original version of 37 CFR §§1.821 – 1.825 effective October 1, 1990 (old rules) and the revised version (new rules) effective July 1, 1998 as well as World Intellectual Property Organization (WIPO) Standard ST.25.
Checker Version 3.0 replaces the previous DOS-based version of Checker, and is Y2K-compliant. Checker allows public users to check sequence listings in Computer Readable form (CRF) before submitting them to the United States Patent and Trademark Office (USPTO). Use of Checker prior to filing the sequence listing is expected to result in fewer errored sequence listings, thus saving time and money.

Checker Version 3.0 can be down loaded from the USPTO website at the following address:
http://www.uspto.gov/web/offices/pac/checker

VERIFICATION SUMMARY DATE: 10/16/2001
PATENT APPLICATION: US/09/966,422 TIME: 15:52:21

Input Set : A:\30534111.app
Output Set: N:\CRF3\10162001\I966422.raw

L:16 M:270 C: Current Application Number differs, Replaced Current Application Number
L:1336 M:258 W: Mandatory Feature missing, <220> FEATURE:
L:1336 M:258 W: Mandatory Feature missing, <223> OTHER INFORMATION:
L:2066 M:258 W: Mandatory Feature missing, <221> not found for SEQ ID#:67
L:2066 M:258 W: Mandatory Feature missing, <222> not found for SEQ ID#:67
L:2066 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:67
L:2067 M:258 W: Mandatory Feature missing, <221> not found for SEQ ID#:67
L:2067 M:258 W: Mandatory Feature missing, <222> not found for SEQ ID#:67
L:2067 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:67
L:2080 M:258 W: Mandatory Feature missing, <221> not found for SEQ ID#:68
L:2080 M:258 W: Mandatory Feature missing, <222> not found for SEQ ID#:68
L:2080 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:68
L:2081 M:258 W: Mandatory Feature missing, <221> not found for SEQ ID#:68
L:2081 M:258 W: Mandatory Feature missing, <222> not found for SEQ ID#:68
L:2081 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:68

<210> 16
<211> 41
<212> PRT
<213> Artificial Sequence

<400> 16

Errored
Actual file contents
as of 11/02/01 mst
12:07 pm

A 213 response of "Artificial Sequence" requires an explanation in Field 223.
MVA

FYI: Sequence 67 and Sequence 68 unknowns must be represented in fields 221, 222 and 223 as "unsure", location and possible values of nucleotide residues.
MVA

258 cggcaatcgc agtgcacatg
261 <210> SEQ ID NO: 6 80
262 <211> LENGTH: 23
263 <212> TYPE: DNA
264 <213> ORGANISM: Artificial Sequence
266 <220> FEATURE:
267 <223> OTHER INFORMATION: Description of Artificial Sequence:Synthetic
268 oligos
270 <400> SEQUENCE: 6
271 cagacaccat taacatcccg aat
274 <210> SEQ ID NO: 7 23
275 <211> LENGTH: 22
276 <212> TYPE: DNA
277 <213> ORGANISM: Artificial Sequence
279 <220> FEATURE:
280 <223> OTHER INFORMATION: Description of Artificial Sequence:Synthetic
281 oligos
283 <400> SEQUENCE: 7
284 agaatgaaat gccgaggaag ag
287 <210> SEQ ID NO: 8 22
288 <211> LENGTH: 1230
289 <212> TYPE: PRT
290 <213> ORGANISM: RAT
292 <400> SEQUENCE: 8
293 Met Cys Pro Pro Gln Leu Phe Ile Leu Met Met Leu Leu Ala Pro Val
294 1 5 10 15
296 Val His Ala Phe Ser Arg Ala Pro Ile Pro Met Ala Val Val Arg Arg
297 20 25 30
299 Glu Leu Ser Cys Glu Ser Tyr Pro Ile Glu Leu Arg Cys Pro Gly Thr
300 35 40 45
302 Asp Val Ile Met Ile Glu Ser Ala Asn Tyr Gly Arg Thr Asp Asp Lys
303 50 55 60
305 Ile Cys Asp Ser Asp Pro Ala Gln Met Glu Asn Ile Arg Cys Tyr Leu
306 65 70 75 80
308 Pro Asp Ala Tyr Lys Ile Met Ser Gln Arg Cys Asn Asn Arg Thr Gln
309 85 90 95
311 Cys Ala Val Val Ala Gly Pro Asp Val Phe Pro Asp Pro Cys Pro Gly
312 100 105 110
314 Thr Tyr Lys Tyr Leu Glu Val Gln Tyr Glu Cys Val Pro Tyr Lys Val
315 115 120 125
317 Glu Gln Lys Val Phe Leu Cys Pro Gly Leu Leu Lys Gly Val Tyr Gln
318 130 135 140
320 Ser Glu His Leu Phe Glu Ser Asp His Gln Ser Gly Ala Trp Cys Lys
321 145 150 155 160
323 Asp Pro Leu Gln Ala Ser Asp Lys Ile Tyr Tyr Met Pro Trp Thr Pro
324 165 170 175
326 Tyr Arg Thr Asp Thr Leu Thr Glu Tyr Ser Ser Lys Asp Asp Phe Ile
327 180 185 190
329 Ala Gly Arg Pro Thr Thr Thr Tyr Lys Leu Pro His Arg Val Asp Gly

RAW SEQUENCE LISTING DATE: 10/16/2001
PATENT APPLICATION: US/09/966,422 TIME: 15:52:20

Input Set : A:\30534111.app
Output Set: N:\CRF3\10162001\I966422.raw

201 gccaaatatt aatgtattta ttagtttatc atatttaatt tttataacac cataaatagg 720
202 tattaatgta cacattttat agatgaggaa aatgtggttc tgagaggtga agcattttgc 780
203 ctagtgatca cagctaaaaa gtgatagagc tgttctttat tttaaagttc acattgtact 840
204 accctggctc cctaatcaca gatgggcagg gtaggggttg ggtggggaca gaagttggag 900
205 agtggatgtg gctgccaacc acacaagttg tgccaaccca cagattgagg aaagatgcta 960
206 aatttggaat ctggcaaacc agtgtttggt tcttagctct gccacttcta agctgtgtga 1020
207 aacttggttg aggtccctaa cttctcctga gggtgaacaa ctcacaaagt tgttttgctt 1080
208 attaaatgtg ataacacctg taaacatcta acagagtgcc tagcacatag cagggatcta 1140
209 gcaattgaat tagggttatt tgtttctgtc tactgattgg gtattgtttc tgacacttac 1200
210 ccaagtgtga atagcctata acactggtat aatttgtgaa atgatgctgc catctagtga 1260
211 aaaccaagac acacacacac acacacacac acacacacat acacacacac gtgcgcgcgc 1320
212 atggacaccc agcttcacca atgacaatat ggattggcat gttttagcct cacaacacag 1380
213 agccctgggg ctaactggca cctagagagg tcatctcggc cagtgccttc caaactacca 1440
214 gtgctgaaaa gccagttcaa aaaattttga acccattgca caccaatatt tttgtgaaat 1500
215 accataaaaa taaattactg gaaaaatgaa ataaaaaata tgtataaaat acaaaccaaa 1560
216 attttagaac tgttagattc aacagcaaaa aattgctgta tacatctctg accaattgct 1620
217 ttcagtttct gtgcttatct ctctacgacc tttgtaacac acagtgaacc agcgctggcc 1680
218 catggataca ctctagtagc cccaatctag ctaaggcagc cccttatagt taatcaatcc 1740
219 tgtcaaacag gaaaggctgg caaaaccact ggtctgcatg tactttgtcc tttacacaag 1800
220 gaaggatgca aacgtggaaa actgagtgga catggtgttc aggagattga ggctcagcta 1860
221 aattccagct tatttacctg cagttgctta caaagtgttt ggacataatt gtgtaaagct 1920
222 agggtttttt ttctggtttt taaaacaggt aaaggatgtc acagcaccac ttaataacat 1980
223 ttcttctgaa gtccagattt taacatctga tgccaataaa ttaactgctg agaacatcac 2040
224 tagtgctacg cgagtggttg gacagatatt caacacttcc agaaatgctt cacctgaggc 2100
225 aaagaaagtt gccatagtaa cagtgagtca actcctagat gccagtgaag atgcttttca 2160
226 aagagttgct gctactgcta atgatgatgc ccttacaacg cttattgagc aa 2212
229 <210> SEQ ID NO: 4
230 <211> LENGTH: 449
231 <212> TYPE: DNA
232 <213> ORGANISM: Homo sapiens
234 <220> FEATURE:
236 <400> SEQUENCE: 4
237 acagtaaaac ttacctgttg tggtcttttt aatcacctcg tttgagtttt atctgtttct 60
238 ctcctttatt tcccagtcct ctcagaaagt cttcctcaat gtattttgct caggattaag 120
239 aattagataa aacctgttgt ttattattat tcggcataat ggacttggta gtttttctat 180
240 ttttcaatag atttgtactt gaataaggtg aagaatttca cacaacatac aagagtacca 240
241 ttgttcctta tatcgttaaa tctttgtgac acactttgac aaaaatgtag aacctataac 300
242 aaattctttt acaagttact ataaaggaca caaagagaaa actttacctt ccagaacaaa 360
243 atgactcctg atgaacagtg tgtggggatt tgcttgtatg tattaaactt ttgacctctg 420
244 aaaaaaaaaa aaaaaaaaaa aaaaaaaag 449
247 <210> SEQ ID NO: 5
248 <211> LENGTH: 80
249 <212> TYPE: DNA
250 <213> ORGANISM: Artificial Sequence
252 <220> FEATURE:
253 <223> OTHER INFORMATION: Description of Artificial Sequence:Synthetic
254 oligos
256 <400> SEQUENCE: 5
257 gctgtgcagc gctgagtgcg ttccaggtaa atgtcactaa cagaaaatag tgcagtaagg 60

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/966,422

DATE: 10/16/2001
TIME: 15:52:20

Input Set : A:\30534111.app
Output Set: N:\CRF3\10162001\I966422.raw

gttgtttttg gaattacctg gattctagca tacctgatgc tagttaatga tgatagcatc 1380
aggatcgtct tcagctacat attctgcctt ttcaacacta cacagggatt gcaaattttt 1440
atcctgtaca ctgttagaac aaaagtcttc cagagtgaag cttccaaagt gttgatgttg 1500
ctatcgtcta ttgggagaag gaagtcattg ccttcagtga cgcggccgag gctgcgtgta 1560
aagatgtata atttcctcag gtcattgcca accttacatg aacgctttag gctactggaa 1620
acctctccga gtactgagga aatcacactc tctgaaagtg acaatgcaaa ggaaagcatc 1680
tag 1683

<210> SEQ ID NO: 2
<211> LENGTH: 560
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens
<400> SEQUENCE: 2

Met Glu Thr Tyr Ser Leu Ser Leu Gly Asn Gln Ser Val Val Glu Pro
1 5 10 15
Asn Ile Ala Ile Gln Ser Ala Asn Phe Ser Ser Glu Asn Ala Val Gly
20 25 30
Pro Ser Asn Val Arg Phe Ser Val Gln Lys Gly Ala Ser Ser Ser Leu
35 40 45
Val Ser Ser Ser Thr Phe Ile His Thr Asn Val Asp Gly Leu Asn Pro
50 55 60
Asp Ala Gln Thr Glu Leu Gln Val Leu Leu Asn Met Thr Lys Asn Tyr
65 70 75 80
Thr Lys Thr Cys Gly Phe Val Val Tyr Gln Asn Asp Lys Leu Phe Gln
85 90 95
Ser Lys Thr Phe Thr Ala Lys Ser Asp Phe Ser Gln Lys Ile Ile Ser
100 105 110
Ser Lys Thr Asp Glu Asn Glu Gln Asp Gln Ser Ala Ser Val Asp Met
115 120 125
Val Phe Ser Pro Lys Tyr Asn Gln Lys Glu Phe Gln Leu Tyr Ser Tyr
130 135 140
Ala Cys Val Tyr Trp Asn Leu Ser Ala Lys Asp Trp Asp Thr Tyr Gly
145 150 155 160
Cys Gln Lys Asp Lys Gly Thr Asp Gly Phe Leu Arg Cys Arg Cys Asn
165 170 175
His Thr Thr Asn Phe Ala Val Leu Met Thr Phe Lys Lys Asp Tyr Gln
180 185 190
Tyr Pro Lys Ser Leu Asp Ile Leu Ser Asn Val Gly Cys Ala Leu Ser
195 200 205
Val Thr Gly Leu Ala Leu Thr Val Ile Phe Gln Ile Val Thr Arg Lys
210 215 220
Val Arg Lys Thr Ser Val Thr Trp Val Leu Val Asn Leu Cys Ile Ser
225 230 235 240
Met Leu Ile Phe Asn Leu Leu Phe Val Phe Gly Ile Glu Asn Ser Asn
245 250 255
Lys Asn Leu Gln Thr Ser Asp Gly Asp Ile Asn Asn Ile Asp Phe Asp
260 265 270
Asn Asn Asp Ile Pro Arg Thr Asp Thr Ile Asn Ile Pro Asn Pro Met
275 280 285
Cys Thr Ala Ile Ala Ala Leu Leu His Tyr Phe Leu Leu Val Thr Phe